12.7. The Forward Theorem

**Theorem 12.11.** Assuming finite source and reproduction alphabets and a discrete memoryless source, if \( R > \tilde{R}(D) \), then \((R, D)\) is achievable.

**Proof.** Let \( p \) be the pmf of the discrete memoryless source. Let \( W \) be such that
\[
E[d(X,Y)] \leq D \quad \text{and} \quad \tilde{R}(D) = I(p \times W).
\]
We show that if \( R > \tilde{R}(D) \), then \((R, D)\) is achievable. Let \( \Delta D > 0 \) and \( \Delta R > 0 \) be given. Choose \( \varepsilon \) so that
\[
0 < \varepsilon < \frac{R - I(p \times W)}{3} \quad \text{and} \quad \varepsilon < \frac{\Delta D}{3}.
\]
With these choices of \( P_{XY} \) and \( \varepsilon \), let \( \hat{A}_n \) be as above with \( n \) so large that
\[
P((X, Y) \notin \hat{A}_n) < \varepsilon/d_{\max} \quad \text{and} \quad e^{-\exp(n[R-I(p\times W)-3\varepsilon])} < \varepsilon/d_{\max},
\]
where
\[
d_{\max} := \max_{x,y} d(x,y).
\]
Let \( N := \lceil \exp(nR) \rceil \). Given \( y_1, \ldots, y_N \) each in \( Y^n \), consider the mapping \( q:X^n \to Y^n \) defined by \( q(x) = y_i \) if \((x, y_i) \in \hat{A}_n\). If there is more than one such \( i \), use the smallest, and if there is no such \( i \), put \( q(x) = y_N \). It is convenient to put
\[
B := \bigcup_{i=1}^{N} B_i, \quad \text{where} \quad B_i := \{x : (x, y_i) \in \hat{A}_n\}.
\]
Then if \( x \in B \), there is an \( i \) with \((x, y_i)\) distortion typical, and thus \( d_n(x, q(x)) \leq D + \varepsilon \). Hence,
\[
E[d_n(X, q(X))] = E[d_n(X, q(X)) I_B(X)] + E[d_n(X, q(X)) I_{B^c}(X)]
\]
\[
\leq D + \varepsilon + d_{\max} P(X \in B^c).
\]
We now employ the random coding argument. Let $Y \sim pW$. To analyze this last probability, write

$$\mathbb{E}(X \in B^c) = \sum_x p^n(x) I_{B^c}(x) = \sum_x p^n(x) \prod_{i=1}^N I_{B_i^c}(x) = \sum_x p^n(x) \prod_{i=1}^N I_{A_i}(x, y_i)$$

$$=: \alpha(y_1, \ldots, y_N).$$

We now employ the random coding argument. Let $Y_1, \ldots, Y_N$ be i.i.d. with common pmf $(pW)^n(y)$. Then

$$\mathbb{E}[\alpha(Y_1, \ldots, Y_N)] = \sum_x p^n(x) \mathbb{E}[I_{A_n}(x, Y)]^N.$$ 

Next write

$$\mathbb{E}[I_{A_n}(x, Y)] = 1 - \mathbb{E}[I_{\hat{A}_n}(x, Y)].$$ 

Then

$$\mathbb{E}[I_{\hat{A}_n}(x, Y)] = \sum_{y: (x, y) \in \hat{A}_n} P^n(y)$$

$$\geq \sum_{y: (x, y) \in \hat{A}_n} P^n_X(y|x) \exp(-n[I(P_{XY}) + 3\varepsilon]), \quad \text{by Theorem 12.10}$$

We now have

$$\mathbb{E}[\alpha(Y_1, \ldots, Y_N)]$$

$$\leq \sum_x p^n_X(x) \left[ 1 - \sum_{y: (x, y) \in \hat{A}_n} P^n_X(y|x) \exp(-n[I(P_{XY}) + 3\varepsilon]) \right]^N$$

$$\leq \sum_x p^n_X(x) \left[ 1 - \sum_{y: (x, y) \in \hat{A}_n} P^n_Y(y|x) + e^{-N \exp(-n[I(P_{XY}) + 3\varepsilon])] \right], \quad \text{by Lemma 12.9}$$

$$= 1 - \sum_{(x, y) \in \hat{A}_n} P^n_{XY}(x, y) + e^{-N \exp(-n[I(X \wedge Y) + 3\varepsilon])}$$

$$= P((X, Y) \in \hat{A}_n^c) + e^{-N \exp(-n[I(X \wedge Y) + 3\varepsilon])}$$

$$\leq P((X, Y) \in \hat{A}_n^c) + e^{-\exp(n[R - I(X \wedge Y) - 3\varepsilon])}$$

$$< 2\varepsilon/d_{\max}.$$ 

Hence, there is a realization of $y_1, \ldots, y_N$ with $\alpha(y_1, \ldots, y_N) < 2\varepsilon/d_{\max}$. Returning to (12.2), we have

$$\mathbb{E}[d_n(X, q(X))] \leq D + 3\varepsilon < D + \Delta D.$$ 

We saw earlier that $D \geq D_{\max} \Rightarrow R(D) = 0$. We also saw that for a DMS with pmf $p$, $D \geq D_{\max} \Rightarrow \bar{R}(D) = 0$. In the case of a DMS, we can find all $D$ for which $\bar{R}(D) = 0$. Of course this is equivalent to finding all $D$ for which $\bar{R}(D) = 0$. 

114
Proposition 12.12. Assuming finite source and reproduction alphabets, a discrete memoryless source with pmf $p$ satisfies $R(D) = 0 \iff D \geq D_*$, where

$$D_* := \min_y \sum_x p(x)d(x,y).$$

Proof. We first show that $D \geq D_* \Rightarrow \tilde{R}(D) = 0$. Let $y_0$ be any value of $y$ achieving the minimum in the definition of $D_*;$ i.e.,

$$\sum_x p(x)d(x,y_0) = D_*.$$

Put

$$W_0(y) := \begin{cases} 1, & y = y_0, \\ 0, & y \neq y_0. \end{cases}$$

If $P_{XY}(x,y) = p(x)W_0(y)$, then $X$ and $Y$ are independent, and so $I(p \times W_0) = 0$. Also,

$$E[d(X,Y)] = \sum_x \sum_y p(x)W_0(y)d(x,y) = \sum_x p(x)d(x,y_0) = D_*.$$

These observations show that for $D \geq D_*$,

$$\tilde{R}(D) = \inf_{E[d(X,Y)] \leq D} I(X \land Y) \leq I(p \times W_0) = 0.$$

To prove that $\tilde{R}(D) = 0 \Rightarrow D \geq D_*$, suppose

$$0 = \tilde{R}(D) = \inf_{W: E[d(X,Y)] \leq D} I(p \times W).$$

Since the infimum must be achieved, there is a $W$ with $E[d(X,Y)] \leq D$ and $I(p \times W) = 0$. Now $I(X \land Y) = 0$ implies $X$ and $Y$ are independent, which means that $W(y|x) = W(y)$ does not depend on $x$. Hence,

$$D \geq E[d(X,Y)] = \sum_x \sum_y p(x)W(y)d(x,y) = \sum_y W(y) \left[ \sum_x p(x)d(x,y) \right] \geq \sum_y W(y)D_* = D_*.$$

$\square$